YALE JOURNAL OF BIOLOGY AND MEDICINE 46, 264–283 (1973)

An Analysis of Diagnostic Reasoning

II. The Strategy of Intermediate Decisions[1]

ALVAN R. FEINSTEIN[2]

Cooperative Studies Support Center and the Department of Medicine of the West Haven Veterans Administration Hospital, and the Departments of Medicine and Epidemiology of the Yale University School of Medicine

Received for publication 28 November 1972

The process of diagnostic reasoning contains a sequence of intellectual "stations" through which a patient's manifestations are transferred en route to the diagnostic entities that emerge as explanatory conclusions. The previous paper of this series (1) contained an account of the domains, disorders, derangements, and other entities that act as "stations" during the directional sequence of the reasoning. In this paper, I shall try to indicate the way in which the transfers through these "stations" are effected by a rational series of intermediate decisions that are commonly subsumed under the name of judgment.

A. CHOICE OF PERTINENT MANIFESTATIONS

Although statistical and computer approaches to diagnosis begin with a collection of "input" data, some of the fundamental decisions of diagnostic reasoning occur long before any data are selected as "input" to the analytic process.

1. Authentication of Evidence

Since an act of perception and description is required for an observed entity to be converted into "data," the clinician's first main decisions involve the authenticity of the basic evidence that constitute the data with which he plans to think.

In the information that constitutes demographic and symptomatic data, a stimulus, sensation, or event had to be perceived by a patient, described by that patient, perceived by a receptor "apparatus" (which is usually a physician), and converted into a descriptive statement. Errors can occur at each step in this procedure: the

[1] Supported by PHS Grant Number HS 00408 from the National Center for Health Services Research and Development.

[2] Professor of Medicine and Epidemiology, Yale University School of Medicine, New Haven, Connecticut; Chief, Clinical Biostatistics and Cooperative Studies Support Center, Veterans Administration Hospital, West Haven, Connecticut.

patient may be unreliable or uncommunicative; the doctor may be imperceptive, failing to ask or note enough; the doctor may be overzealous, using leading questions or badgering questions to maintain the preconceptions that elicit a "typical" but inaccurate history; or an inanimate recipient of information, such as a printed form or a computer, may be confusing to the patient or inadequate for the expression of subtle details.

In the physical examination, a sign is noted when certain sensory stimuli have passed from the observed entity to the doctor's cerebral recognition. Errors can occur before these stimuli reach the doctor's mind because of sensory defects in his vision, hearing, or palpation, or because he may fail to examine certain regions. Alternatively, despite satisfactory sensory receptors and apparently complete examination, the doctor's cerebration may still fail to recognize and identify such entities as a faint aortic diastolic murmur or peripheral retinopathy.

In paraclinical data, the opportunities for error are enhanced by the many intermediate steps that occur as a substance derived from a patient is transported to its locale of examination, processed there, and eventually transformed into a report in the patient's record. The substance may have been taken from the wrong patient, given a wrong label, or sent to the wrong laboratory. In the laboratory, or other locale of examination, defects may have existed in the test itself, in the apparatus, in the reagents, in the criteria for appraisal, or in the human appraiser. The reported result may have been distorted as it underwent various transcriptions before reaching the patient's record.

At this elemental stage of data gathering, therefore, a careful clinician will arrange to verify the basic facts before any other reasoning begins.

2. Designation of Evidence

Although symptoms and signs are sometimes regarded as the fundamental elements of clinical data (2), these entities are actually "diagnoses" that represent labels or designations given to the basic observations that were just described. For example, a patient seldom reports the symptom of pleuritic pain; the patient may say that he gets a stabbing sensation in his chest when he takes a deep breath, and the doctor then designates this event with the symptom name of pleuritic pain. Similarly, the "signs" of petechiae, apical holosystolic murmur, and basilar dullness are interpretations of the fundamental observations or sensations encountered by the doctor during physical examination.

Many of the problems in making these decisions of designation or interpretation have been discussed elsewhere (3). The sources of error include: for symptoms, the failure of the doctor and patient to communicate in meaningful or precise terms (e.g., what does a patient or doctor mean by gas pain, dyspepsia, or heartburn?); and, for signs, the doctor's inaccuracy in interpreting the sensations (e.g., calling a systolic murmur diastolic) or the doctor's neglect of alternative causes for the observed evidence (e.g., designating a palpable rectus muscle as liver, or a pulmonary wheeze as a pleural rub).

3. Contemplation of Deviance

After the observed evidence has been designated, the clinician's next step is to decide which entities are sufficiently deviant from the accepted range to be called "abnormal." These decisions involve the answers to such questions as: when is a palpable liver not enlarged; which systolic murmurs are physiologic; how much

shortness of breath is dyspnea; how high is tachypnea, hypertension, or leukocytosis and how low is bradycardia, oliguria, anemia, or thrombocytopenia?

The problems of establishing an appropriate range of normal for clinical and paraclinical data are beyond the scope of this discussion, but the main point to be noted here is that such ranges have not been established in a scientifically satisfactory manner for many critical topics in medical data (3). A prime source of difficulty is not in clinical observation or in laboratory technology, but in epidemiologic sampling. The populations used for assessing the "normal range" of most laboratory tests have been inadequate in their range of age, race, sex, ethnic groups, and other demographic features. In addition, a major controversy exists about whether the idea of "normal" should be decided according to biologic health or statistical calculation (4, 5). For the calculation, statisticians still debate whether "abnormal" should be measured according to percentiles of a distribution, or distance from a mean (6, 7); and clinicians are uncertain about either form of measurement in the absence of satisfactory long-term follow-up data to demonstrate that the measured "abnormality" in statistics represents a pathologic outcome in biologic "disease."

4. Consideration of Pertinence

After all these decisions have been made about the assembled information, the clinician then determines which normal or abnormal findings will be ignored, and which ones will receive further attention. Among the abnormal findings that might be ignored are an asymptomatic ingrown toenail or an ocular cataract that has produced no visual symptoms. Among the normal findings that might receive major attention are "pertinent negatives" used to support or refute certain other evidence. Thus, the maintenance of customary weight is a "normal" but contradictory finding in a patient who claims to have had profound anorexia for a long period of time; and the absence of appropriate exposure suggests a cause other than psittacosis in a patient with viral pneumonia.

Certain findings may have therapeutic pertinence even though the findings are either normal, or not subjected to diagnostic investigation. For example, a structurally and functionally normal nose or breast may be esthetically unsatisfactory to the patient, and may thereby be subjected to plastic surgery. A hospitalized patient may be treated for constipation, insomnia, or anxiety although these symptomatic discomforts may not receive further diagnostic assessment. In certain urgent circumstances, such as severe gastrointestinal hemorrhage or acute pulmonary edema, appropriate ad hoc treatment may be given before or while the diagnostic procedures continue.

After the pertinent findings have been selected, the clinician has assembled the list of manifestations with which "formal" diagnostic reasoning can begin.

These four activities may seem so obvious to a well-functioning clinician that he may be impatient with their detailed discussion. Nevertheless, perhaps because they seem so obvious, the prediagnostic decisions are constantly overlooked both during the training of medical students and during the composition of computer programs that attempt to achieve automated diagnosis. The medical student often fails to develop a rigorous rational discipline of the bedside because he is taught to think precisely about the "basic" morphologic, physiologic, biochemical, or molecular explanations for clinical phenomena, but he may receive no comparable intellectual instruction for the challenge of determining the basic authenticity, desig-

nation, deviance, and pertinence of the elemental data that identify the explained phenomena. In the absence of such a specific scientific discipline, the student learns to make the decisions by intuitive, judgmental, or other nondescript methods that he later, as a consultant physician, cannot specify when he tries to describe the diagnostic process to medical students, statisticians, or computer programmers. Consequently, the computer or card-sorting programs developed for history-taking, and the statistical procedures developed for diagnostic "reasoning," may be grossly distorted and oversimplified, producing defective data and defective reasoning.

B. THE PROCESS OF SUBSEQUENT SELECTIONS

As diagnostic reasoning proceeds in the search for the "immediate cause" (8) that connects one phenomenon to the next, the clinician will constantly evaluate sets of candidates from which to make choices. The process of including and subsequently excluding candidates in such sets is often described with the phrase "rule out."

After all the appropriate candidates are assembled in the set, the merits of the first candidate are evaluated, and the candidate is either kept in the set or "ruled out." The next candidate then receives similar considerations, and the process continues until all candidates have been evaluated. If more than one candidate remains after this first appraisal of the entire set, the "survivors" are now reevaluated both in relation to the role for which they compete, and in relation to one another. The iteration of these appraisals continues until all candidates but one have been excluded. The one that remains is selected.

The selection process for each decision in diagnostic reasoning thus contains at least three distinct components: a set of candidates; a mechanism for "nominating" or including candidates in the set; and a mechanism for excluding or "ruling out" all but the ultimate choice.

1. Formation of Sets

The set assembled for each decision will vary according to the entity about which the decision is being made. Some of the sets will be topographic, containing a list of anatomic structures or physiologic functions. Other sets will be explicatory, containing lists of disorders, derangements, and other abnormalities that are regarded as "immediate causes" for the antecedent entity.

Since a manifestation must be explained by citation of both a domain and a disorder, at least two different sets will be assembled for each manifestation. As the domains are considered, a third set of candidates may be assembled to select a focus within each domain. For example, the topographic set for the manifestation, epigastric pain, will consist of all domains that might be the source of this manifestation. In this case, the candidate domains would be the chest and the abdomen. The topographic set for each of these domains would consist of all focal structures (heart, esophagus, pleura, etc., and stomach, duodenum, liver, pancreas, etc.) that compose the domain. The explicatory sets for the manifestation would consist of all the disorders that can occur at these focal structures. Such disorders in this case would include muscular spasm, mucosal irritation, etc.

A set can be named according to the entity to which it refers. Depending on the way the sets are used, the names can refer either to the contents of the set, or to the entity explained by the set. A domain set contains its constituent foci. Thus, the domain set of the gastrointestinal channel contains all the structures ex-

tending from the mouth to the anus through such intermediaries as stomach and intestines. Examples of other domain sets were discussed previously (1) when domains and foci were initially defined. A primary disorder set consists of all the disorders than can explain a manifestation in a particular domain; and the set is usually named after the manifestation. Thus, the tachypnea set would contain all disorders that can produce tachypnea. A secondary disorder set consists of all the disorders or derangements that can cause a primary disorder, and the set is named according to the primary disorder. Thus, the cardiomegaly set would contain such disorders as acute cardiac dilatation and such derangements as rheumatic heart disease.

2. Implications of Manifestations

The first major step in diagnostic reasoning is to assemble the sets of domains, foci, and disorders that can account for a manifestation. Because of preclinical training in gross anatomy, in histopathology, and in those aspects of physiology, biochemistry, and microbiology that remain important for explaining gross disorders of structure and function, a clinician can usually immediately choose the domains that are implied by a patient's manifestations.

Certain manifestations will regularly imply an affected organ; thus, the eye is implicated by ocular pain; the larynx, by hoarseness; and the bladder, by dysuria. Certain manifestations point to a region; such symptoms as pain in the head, neck, chest, abdomen, or back, and such physical signs as lower thoracic percussive dullness, precordial thrills, distention of abdomen, and swelling of a leg all imply a problem in the associated regions. The involvement of a channel is suggested by such symptoms as cough, vomiting, dysmenorrhea, and diarrhea, and by such signs as distended neck veins, absent arterial pulsations, and abnormal appearances of sputum, urine, feces, etc. Other symptoms or signs are referred to systems. Delirium (sensorial system), hypertension (cardiovascular system), dyspnea (cardiopulmonary–vascular–hemic system), fever (body as a whole), and edema (cardiovascular–metabolic system) are examples of manifestations and of their associated systems.

3. Inclusion and Exclusion of Candidates

As described elsewhere (9), some of the main errors of diagnostic reasoning arise during the processes of including and excluding candidates. With errors of omission, the diagnostician does not include enough candidates when he assembles his list. With errors of commission, his exclusions are defective, leading to a wrong choice or to an unnecessary choice.

Some of the errors occur late in the reasoning process, as the clinician nears the goal of a diagnostic entity, but many others occur early, during the evaluation of manifestations. Among the most common types of error are those that occur when the clinician makes a peremptory decision, leaping to the choice of a particular candidate without contemplating or giving adequate weight to all the alternative possibilities.

a. Errors of omission. An error of omission in the diagnostic analysis of a channel often occurs when a manifestation at one end of the channel is ascribed to a nearby structure in the channel, without consideration of the possibility that the true source is more distant. Thus, the signs and symptoms of peripheral arterial disease may be attributed to disease of small vessels because the peripheral pulses

are not palpable, but the actual difficulty may lie higher in the arterial channel, at the aortic bifurcation. Similarly, rectal bleeding may be attributed to hemorrhoids or a rectal lesion, without contemplation of diverticula or of lesions located higher in the colon or intestine. A stroke may be ascribed to an intracerebral vascular difficulty, and a lesion of the carotid artery may be overlooked.

In dealing with a region, the diagnostician may neglect to consider all the regional constituents and may concentrate instead only on those that are commonly associated with the observed manifestation. For example, a rub or murmur occurring in the region of the anterior chest may be immediately assigned to a focus in the heart, and the pleura or great vessels may be neglected as possible sources.

When the abnormal function of a system is appraised, the diagnostician may not include all of the physiologic elements that normally enter the activity of the system. For example, the role of metabolic protein products in maintaining the homeostasis of extracellular fluid may be ignored in a patient who has edema. Consequently, the edema may be attributed to a disorder in the cardiovascular system, and a hypoproteinemic cause may be overlooked.

b. Errors of disorientation. Errors of this type occur when the clinician, regardless of completeness in his list of candidates, becomes oriented toward the wrong candidate.

The choice of a region may be disoriented when the diagnostician forgets about the previously (1) described diaphragmatic distortions, channel transmissions, and neurologic peculiarities that can cause a manifestation to appear at a site remote from its provocative source. In these circumstances, to account for the observed manifestation, the diagnostician may need to identify at least two different sequential disorders in two different domains. For example, to account for the manifestation of hoarseness, the first explanatory disorder might be a paralysis of vocal cord; the next explanatory disorder would then be moved to a different location and listed as intramediastinal compression of the recurrent laryngeal nerve.

Without a constant awareness of such patterns of referral, the diagnostician observing a manifestation in a particular region may restrict his focus to organs in that region, and may fail to contemplate a more remote region that may be the true source of the difficulty. In this way, ascites may be ascribed to a hepatic disorder although caused by congestive heart failure; scapular pain due to gallstones may receive orthopedic attention; the pain of a myocardial infarction may be attributed to indigestion; and the manifestations of Horner's syndrome, caused by apical lung cancer, may be "worked-up" for a brain tumor.

Even when the proper region has been selected and all its constituents considered, the investigator may choose the wrong focus within the region because he neglects the capacity of the correct focus to cause the manifestation. For example, when a patient has inspiratory pain in the chest, the diagnostician may consider such foci as rib, pleura, and pericardium, and may exclude the muscles of the chest wall because he may forget that a strained muscle can produce inspiratory pain.

C. THE CORRELATION OF STRUCTURE, FUNCTION, AND EXPLANATION

Although inspection, palpation, roentgenography, and diverse paraclinical tests may demonstrate an anatomic lesion, structural evidence can never alone provide explanations for the functional phenomena that occur as symptoms. Thus, the findings of cardiomegaly or narrowed coronary vessels may not account for chest pain;

the demonstration of gallstones does not necessarily explain abdominal pain; osteoarthritis may not be the cause of back pain; and an abnormal frontal sinus need not account for headache.

Even when the morphologic evidence shows the actual lesion that produces the symptoms of a functional disorder, a mere citation of the lesion does not explain the functional process by which the symptom is produced. For example, an anatomic narrowing of the mitral valve does not account for enlargement of the left atrium unless the obstruction has been hemodynamically significant; a stone in the common bile duct does not usually account for complete occlusion of the duct unless the irritated mucosa has become inflamed or edematous.

In addition to correlating structure and function, the diagnostician must further determine whether a well-correlated lesion and dysfunction provide satisfactory explanations for an observed clinical manifestation. Thus, a stenotic mitral valve producing significant hemodynamic obstruction may cause enlargement of the left atrium, but none of these correlated lesions and dysfunctions of the heart accounts for the manifestation of dyspnea unless additional abnormalities are invoked in the lungs or elsewhere.

The obligation to correlate and to explain is one of the most crucial aspects of clinical reasoning, because the correlated explanations are often more important for therapeutic decisions than in choices of diagnostic names. Thus, the clinician may make an accurate diagnosis of gallstones, but if the diagnosed gallstones do not account for the abdominal pain, a cholecystectomy will not solve the patient's problem. Similarly, the clinician may correctly diagnose and correlate mitral stenosis, left atrial enlargement, and atrial fibrillation, but if these disorders do not account for a patient's fatigue, cardiac surgery might be a futile procedure.

The necessity for correlated explanation as a prelude to therapeutic decisions is one of the main distinctions between diagnostic reasoning in a living patient and diagnostic reasoning at a clinicopathologic conference (CPC). Suppose the patient chosen for CPC had had the manifestations of jaundice, ascites, and bloody stools. If the clinical discusser at the CPC concludes that the patient had both hepatic cirrhosis and rectal carcinoma, and if the pathologist later confirms that both diseases were present, the clinician leaves the conference a "winner." In a living patient, however, these diagnostic identifications would have been inadequate, since they do not alone provide enough explanation of the manifestations to enable a choice of therapy. In the patient just cited, the hepatic cirrhosis may have been quiescent, with all of the manifestations caused by metastatic cancer of the rectum; conversely, the cirrhotic liver may have been decompensated and responsible for all of the observed manifestations, with the rectal carcinoma having been localized, asymptomatic, and discovered by "accident" during the routine physical examination. Thus, in two patients with exactly the same major manifestations and exactly the same disease diagnoses, the estimation of prognosis and choice of treatment would have been greatly affected by the decisions made during the correlated explanation of disorders (10).

The correlation of form, function, and explanation involve a subtle interplay of data and reasoning for the different domains of clinical macrobiology.

1. Regions

In dealing with functional manifestations in a region, the diagnostician must consider all the foci within that region, and must decide which ones can produce the

phenomenon that has been observed. With this type of reasoning, augmented by additional data from the history or physical examination, certain structures and explanations can be included or excluded from consideration.

For example, suppose the manifestation to be explained is pain in the chest. The pain-producing structures in this region are skin, subcutaneous tissue, muscles, nerves, bones, pleura, pericardium, myocardium, esophagus, aorta, pulmonary artery, or the upper parts of the splenic and hepatic capsules. A worsening of the pain with inspiration would tend to rule out skin, subcutaneous tissue, nerves, myocardium, aorta, and pulmonary artery. An onset of the pain after trauma, and an appropriately localized anatomic site, suggest muscle or rib. If pressure on the rib evokes no pain, and if isometric contraction of the suspected muscle does evoke pain, then the muscle appears to be the source of the pain, regardless of what the roentgenogram may show.

In dealing with a topographic type of regional manifestation, the diagnostician may not always be able to rely on the availability of roentgenograms, or the roengenographic findings may not explain the observed manifestations. Accordingly, the reasoning should contain a careful anatomic evaluation of the region.

Suppose the manifestation to be explained is unilateral basilar thoracic dullness. The first step in the reasoning is not a leap to a morphologic explanation in the lungs, but a realization that dullness implies a failure of the resonant thoracic response on percussion. Consequently, the first step is to rule out extrapulmonic causes for this abnormal response. The percussion technique may have been poor; a unilateral tumor of the chest wall may block the transmission of the thump; or the intrathoracic space ordinarily occupied by aerated lung may be usurped by upward displacement of liver (on the right side) or by massive cardiac enlargement (on the left). The next step, proceeding inward from extrapulmonic sources, is to rule out pleural causes. If the transmission of resonance is blocked by a pleural effusion, the transmission of breath sounds or fremitus should also be blocked. Thus, if the breath sounds and fremitus are unimpaired, an effusion is unlikely. Finally, the reasoning reaches the pulmonary parenchyma. Is air failing to enter the parenchyma, due to a block in the tracheobronchial channel, or failing to disseminate, due to consolidation caused by atelectasis, inflammation, or tumor? If consolidation is suspected, appropriate physical findings in breath sounds and fremitus can serve to confirm the decision.

2. Channels

The anatomic continuity of a channel simplifies the list of structures to be considered as the focus for a manifestation. At the same time, however, modern opportunities to inspect channels endoscopically or with radiographic opacification have made clinicians rely heavily on these paraclinical tests for demonstrating localization of lesions.

For the gastrointestinal channel, in particular, diagnosticians will often order supplementary paraclinical tests immediately after the reasoning has reached the first disorder "station." For example, the manifestation of hematemesis implies bleeding from the upper GI channel above the ligament of Treitz. (A swallowed hemoptysis or nasooropharyngeal bleeding should have been first ruled out by history-taking and physical examination before the manifestation was designated as hematemesis). The possible foci in the upper GI channel are the esophagus, stomach, duodenum, and jejunum. Since the hematemesis alone has no characteristics that can localize it, the reasoning now shifts to a search for the derangement that might be causing a disorder at the candidate sites. Cutaneous or palpatory manifestations of hepatic cirrhosis would provide supportive evidence for a lesion in the esophagus; a history of postprandial pain relieved by alkali would support the stomach or duodenum as the site of the bleeding; and the existence of a postoperative dumping syndrome is consistent with a bleeding lesion in the jejunum or gastrojejunal stoma—but none of these contemplations can produce an unequivocal conclusion. For precise anatomic diagnosis, endoscopy or roentgenography, or both, will be employed to demonstrate the location of the lesion.

3. Systems

Because many different systems can be the cause of a single manifestation, the diagnostic "review of systems" may become an intricate exercise in reasoning. In particular, before a decision is reached at the disorder "station," the diagnostician may have conducted many excursions into derangement "stations," searching explanations for the candidate disorders.

Consider an example of reasoning for the manifestation of dyspnea. Dyspnea is not likely to be due to the disorder of anemia if the patient shows no pallor to confirm the anemia; and anemia can be ruled out by a paraclinical test. If the patient is cyanotic, but has no auscultatory evidence of congenital heart disease, the dyspnea is probably not due to a cardiac disorder. A lung disorder becomes likely if the dyspnea is better when the patient reclines rather than sits. In the absence of a history or other findings suggestive of a pulmonary derangement, the dyspnea is probably due to a cardiac disorder.

In a patient with fever, the first domain "station" is the body as a whole. As a specific focus is then sought within that domain, the derangement of inflammation is initially sought in the common locations of respiratory and urinary channels. The next most common locations, such as abdomen and heart, might then be considered. If no inflammatory focus is found, the consideration may shift to sources of reaction for the body as a whole. Such sources might include a nonspecific viral infection, a drug reaction, or both.

D. CLUSTERING

Clustering, a common procedure in diagnostic reasoning, occurs as the result of a decision to combine several different entities and to seek a common name or explanation for them, rather than to deal with each individually.

Such combinatorial tactics have been used for centuries in medical nosology to create the diagnostic entities called syndromes. As defined by Durham (11), a syndrome is "a concurrence or running together of constant patterns of abnormal signs and symptoms." Many syndromes have, therefore, been created as acts of descriptive rather than explanatory diagnosis, and the use of a single name for the syndrome has been an act of linguistic convenience to provide an "abbreviated" label for the concordance of multiple entities.

Despite occasional claims (12) that "the syndrome has as its philosophic basis, not specific disease factors but a chain of physiological processes," syndromes have been collected and christened in diverse, inconsistent ways. The name may be based on an aggregate of observed clinical manifestations (*oculocerebrorenal syndrome*), or on the most prominent of the overt manifestations (*dumping syndrome; stiff-man syndrome*). Whenever a specific underlying lesion can be identified, the nomenclature may reflect the location of that lesion (*carpel tunnel syndrome; superior mediastinal syndrome*). Other names are derived from biochemical or metabolic abnormalities (*hemolyticuremic syndrome; syndrome of inappropriate vasopressin secretion*) or from such "etiologic" entities as a drug (*thalidomide syndrome*), an operation (*postpericardiotomy syndrome*), an accident (*crush syndrome*), a hobby (*motorcycle syndrome*), or the site of a gastronomic adventure (*Chinese restaurant syndrome*). The name "syndrome" has even been applied to an isolated radiographic finding (*empty sella syndrome*) that has no associated clinical or functional abnormalities.

When a simple designation cannot provide an appropriate description of the phenomena, an eponym is generally used. Medicine is rich in these eponymic citations, which usually commemorate the physicians who first described the syndrome (*Klinefelter syndrome; Felty syndrome*), but in recent years, the syndrome may be named after the patients who were its victims (*Mast syndrome*), or literary characters who bore resembling features (*Munchausen syndrome; Harlequin syndrome; Pickwickian syndrome*) (13).

Because the term *syndrome* is so nonspecific and has many connotations other than the ones I should like to discuss here, I shall use the word *cluster* for a group of entities that have been combined for a common diagnostic name or explanation.

1. Types of Clusters

a. Dependent. A dependent cluster is composed of several entities that are directly connected in a cause-and-effect sequence, with each entity having presumably led to the next. Examples of such dependent clusters are the sequence of anorexia → weight loss → fatigue, which commonly occurs in cancer or chronic infections. Another example is polyuria → polydipsia in patients with uncontrolled glycosuria, or polydipsia → polyuria in patients with ingestion of excess fluid.

b. Independent. An independent cluster comprises entities that have not led to each other, and that presumptively have a common underlying cause. Examples of such clusters are the following: dyspnea and edema, due to cardiac decompensation; amenorrhea and an abdominal mass, due to a tumor of the reproductive system; and bradycardia and episodes of unconsciousness, due to the unstable heart block that is called the Adams-Stokes syndrome.

2. The Clustering "Station"

The diagnostic "station" at which a group of entities are clustered can be a disorder, a derangement, or a pathoanatomic entity. In the examples cited in the preceding section, all of the manifestations were clustered as disorders. A group of disorders may be combined at a clustering "station" that is usually a derangement, or occasionally, a pathoanatomic entity. For example, the disorders of cardiomegaly, congestive heart failure, and atrial fibrillation could be clustered with the derangement *rheumatic heart disease,* or with the pathoanatomic entity, *mitral stenosis.*

3. Monopathic vs Polypathic Reasoning

Diagnostic reasoning can be monopathic or polypathic in the attempt to assign one cause or many causes as explanations for the observed phenomena. Although the scientific principle of Occam's razor offers an intellectual parsimony that makes monopathic explanations appealing, and although many clinicopathologic conferences are directed toward identification of a single disease, clinical reality contains many polypathic situations. In contrast to the idealized monopathy that may be chosen for educational instruction, polypathy is often the rule rather than the exception in the work of a practicing clinician.

The question of whether, what, and when to cluster is an intricate part of intermediate judgment in the diagnostic process, and a detailed analysis of cluster decisions is beyond the scope of this discussion. In general, the decisions are affected by the candidates available at the next "station" of the reasoning, beyond the "station" at which the cluster is being considered. Thus, in one of the earlier examples, when dyspnea and edema were considered for clustering into the disorder of cardiac decompensation, a prerequisite to the cluster was the existence of a derangement that would account for this disorder. If no such derangement were available to serve as an "immediate cause" for the cardiac decompensation, some other cluster and a suitable causal derangement might be sought. If no such derangement could be found, a cluster at this diagnostic "station" would be untenable, and a polypathic explanation would be sought for the manifestations.

The type of reasoning employed in a cluster decision is further illustrated in the example that follows. To simplify the example, the initial manifestations have been explained by a mixure of disorders, derangements, and other diagnostic entities.

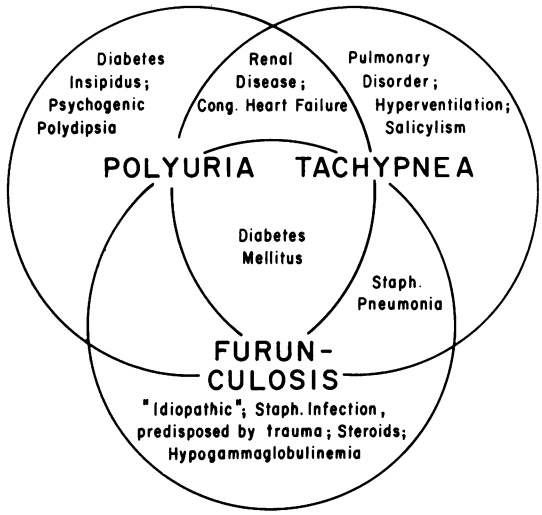

FIG. 1. Venn diagram showing monopathic and polypathic diagnostic choices in the intersection of sets for the explanation of three manifestations. (For further details, see text.)

Suppose a patient's presenting manifestations are polyuria, tachypnea, and furuncles. We could begin by contemplating a set of possible explanations for each of the individual manifestations. For polyuria, the possible causes would include a renal derangement, diabetes insipidus, diabetes mellitus, psychogenic polydipsia, or congestive heart failure (in a phase of diuresis). For tachypnea, the explanations could be congestive heart failure, the hyperventilation syndrome, a pulmonary disorder, a specific pathoanatomic pulmonary entity such as staphylococcal pneumonia, or acidosis due either to excessive salicylate ingestion, to a renal derangement, or to diabetes mellitus. The furuncles could be caused by "idiopathic" staphylococcal infections, or by the patient's cutaneous predisposition due to external trauma, steroid therapy, diabetic hyperglycemia, or hypogammaglobulinemia.

The Venn diagram of Fig. 1 shows the "intersection" of these three sets of explanations and the common distributions for each explanation among the three sets. As noted in the figure, some of the explanations are associated with only one manifestation and others account for two manifestations; but only one—diabetes mellitus—accounts for all three. The monopathic choice in this situation would, therefore, be diabetes mellitus.

This choice would be tempting for a diagnostician who considers only the evidence that has been cited thus far. In clinical reality, of course, additional information would be available to describe the patient's past history, the sequence and timing of events in the present illness, the intake of medications, and the results of a test for glycosuria. A thoughtful diagnostician using this additional evidence might decide that various polypathic diagnoses explain the clinical situation better than the monopathic diabetes mellitus. The diagnosis of diabetes mellitus becomes almost immediately untenable if a suitable test shows no hyperglycemia. As alternative explanations, in appropriate circumstances, a patient might have a renal disease (producing acidosis) or a cardiac disease (producing congestive heart failure) to explain the polyuria and tachypnea, and his occupation might have exposed him to solvents or other agents that led to the furunculosis; during steroid treatment for diabetes insipidus, a patient may have developed a cutaneous infection that led to a staphylococcal pneumonia; finally, a patient with hypogammaglobulinemia, psychically traumatized by his many medical ailments and treatments, might have developed the hyperventilation syndrome and psychogenic polydipsia.

E. LOGICAL VS STATISTICAL REASONING

In the types of judgment that have just been described, a decision is reached at each intellectual "station" in the reasoning. The result of the decision provides

an appellatory name or an explanatory cause for a preceding entity or entities. As background for each decision, the diagnostician assembles an array of data or concepts pertinent to the issue at hand. He then "processes" or evaluates this information by using various strategies to reach a conclusion that represents the decision. These strategies are based on intellectual mechanisms that can be logical or statistical or both.

1. Logical Mechanisms

The term "logical" refers to a mechanism of reasoning in which a single decision is accomplished by being decomposed into a series of sequential decisions, each of which leads successively to another intermediate decision, and ultimately to the conclusion. If the main decision is to choose the disease responsible for a particular manifestation, a logical scheme of reasoning would decompose this decision into successive decisions about manifestations, domains, disorders, derangements, and so on until the responsible disease is cited in the final conclusion.

At each step in these procedures, the logician would like to attain complete certainty for the associated decision. This complete certainty means that a candidate conclusion is either totally excluded (or "ruled out") from consideration, or selected uniquely as the result. (In statistical terms the logician searches for a probability value of either 0 or 1.) This type of logical reasoning requires that all possible candidates be contemplated, that a group of "inclusions" be formed by eliminating the "exclusions," and that subsequent decisions be made to reduce the number of included candidates if a unique choice has not yet been obtained. Thus, if decision A does not yield a unique selection from its group of "inclusions," decisions B, C, and other branching decisions may be performed in an effort to eliminate all but one member of the group. If a unique member cannot be isolated, the logician may then rank the remaining "inclusion" candidates according to qualitative degrees of likelihood.

For example, let us consider diagnostic reasoning for a patient with a suprapubic abdominal mass. The uterus and ovary can be excluded as the source of the mass if the patient is male. Pregnancy is excluded if the patient is female, but above or below the age in which reproduction can occur. Intestinal obstruction is excluded if the bowel movements have been regular. Beyond these exclusions of possible candidates, the remaining reasoning leads to likelihoods and unlikelihoods. Distention of the bladder is unlikely if the patient is a young man without urinary symptoms. Pregnancy is unlikely if the patient has had regular and recent menses. A tumor or cyst is likely if the mass is hard and painless; a cancer is likely if the mass is fixed.

At this stage of reasoning, and with the available information that has just been listed, the diagnostician might not be able to go beyond the cited decisions. His next step, therefore, might be to obtain additional data from such procedures as rectal examination, bimanual pelvic examination, catheterization of bladder, and sigmoidoscopy. If these procedures are inconclusive, the next step might be to order suitable roentgenograms.

The foregoing description has illustrated two additional aspects of logical mechanisms in diagnostic reasoning: (a) each intermediate conclusion was accompanied by a premise containing data specific to that decision and (b) the decisions that could not be expressed with certainty were followed by a plan of action. (In this instance, the planned action was to obtain additional data.) On the basis of these

additional characteristics, we can now outline the logical process of diagnosis according to four basic constituents: a "dissection" of the main decision into a sequence of intermediate decisions; the application of background data or concepts specific to each intermediate decision; the search for certainty in each conclusion; and the development of a plan of action for conclusions that are uncertain.

2. Statistical Mechanisms

In statistical mechanisms of diagnostic reasoning, the original decision is approached immediately, without "dissection" into components. The pertinent substantiating data are assembled directly, and the decision is made on the basis of mathematical strategy and calculation. In this strategy, the background information is cited arithmetically; the numbers are arranged into ratios from which "rates" (or values of probability) are calculated; and the decision is based on the particular candidate that gave either the highest or lowest values in the array of probabilities. These probabilities can be calculated in a direct or indirect manner.

a. Direct probabilities. Let us reconsider the suprapubic abdominal mass cited in the previous section. From results found in other patients with a suprapubic abdominal mass, the diseases associated with the mass might have the following statistical probabilities: 1/10 for distended bladder; 1/350 for carcinoma of bladder; 1/15 for fibroid; 1/20 for pregnancy; 1/150 for carcinoma of cervix; 1/200 for carcinoma of uterus; 1/420 for ovarian cyst; 1/12 for carcinoma of rectosigmoid; 1/25 for carcinoma of ascending colon; etc. On the basis of these probabilities the most likely immediate choice would be a distended bladder, and the least likely would be an ovarian cyst.

Each of the cited calculations is called a "direct" probability, because it is based on direct data for both the numerator and denominator of the calculated rate. Each denominator was based on an observed, counted group of people who manifested a suprapubic mass. Each of the numerators represented the number of those people who had the corresponding disease. The result of the calculation was expressed as a rate of disease per manifestation, and the input data were based on direct information about that rate.

b. Indirect probabilities. The type of rate just cited cannot always be calculated as a direct probability because the necessary information is seldom available. Clinicians often study a population of people with the same disease and calculate the frequency rate for each manifestation of that disease, but the reverse procedure is seldom done. Thus, we can readily find information about the frequency of hemoptysis in patients with lung cancer, but few or no data exist about the frequency of lung cancer in patients with hemoptysis. Another type of readily available information is collected by epidemiologists, and provides the mortality rates of a disease in the general population. In these rates, the denominators are taken from data of the census bureau, and the numerators are assembled from the data entered nationally on death certificates. If we make the unwarranted assumption (14) that these rates of diagnosis at death represent the true rates of disease, we then have two types of information readily available: the rate of a manifestation per disease, and the rate of that disease in the general population. Let us now make the further (and also unwarranted) assumption that a third type of information is available: the rate of the manifestation in the general population.

If the necessary assumptions be granted and the stated rates accepted, a statistical strategy called "Bayes theorem" will provide an indirect method for converting

those three rates into the desired result of a rate of disease per manifestation. According to the Bayesian formula, the rate of disease per manifestation is obtained by multiplying the rate of manifestation per disease times the rate of the disease in the general population; and by then dividing this product by the rate of the manifestation in the general population. Suppose, for example, that a suprapubic abdominal mass occurred in 1/10 of patients with cancer of the rectosigmoid, and in 1/100,000 of patients in the general population; and suppose further that cancer of the rectosigmoid occurred in 1/120,000 of patients in the general population. The Bayesian result for the probability of rectosigmoid cancer in a patient with a suprapubic mass would then be

$$\frac{\frac{1}{10} \times \frac{1}{120{,}000}}{\frac{1}{100{,}000}} = \frac{1}{12}$$

(The numbers here were deliberately contrived to yield the same result stated earlier for the direct probability of this event).

The basic statistical validity of the Bayesian strategy has often been criticized (15–18), but the main scientific issue to be noted here is the unreliability of the data used in the calculations. The rates of diseases and of manifestations in the general population are not determined accurately enough for scientific credence to be placed in their values (14). To overcome this objection to data from the general population, Bayes-ophilic diagnosticians have used a population that can be accurately observed and counted: the patients at hospitals. With this populational transfer, the numbers used as the denominator and as one numerator in the Bayesian calculation are the rates of manifestations and of diseases in the population of all people examined in a hospital, or in the patients seen by a particular service of that hospital.

This choice of population provides reliable data, but the results cannot be applied beyond the confines of that hospital or of that service. As changes occur in nomenclature and concepts of disease, in tactics of "screening" for disease, in accuracy and availability of diagnostic tests, and in patterns of patient referrals, the rates of manifestations and of diseases will fluctuate enormously in any population of hospitalized patients. The "established" rates of manifestation per disease will be altered accordingly (3, 14). Moreover, for a particular service at a hospital, the rate of a manifestation per disease or per patient, and the rate of a disease per patient will rise or fall according to whether the service is medical or surgical, general or specialized, and ward or private (3). In one of the few instances in which an apparently "successful" Bayesian diagnostic procedure was tested at a hospital (19) other than the one at which the procedure was developed (20), the "success" was not confirmed.

To ascertain that the clinical population has not changed in any of the critical variables just cited, Bayesian diagnosticians can regularly assemble data for the three rates that are necessary for the calculations: manifestation per disease, manifestation per patient, and disease per patient. If these rates remain unaltered with the passage of time, the values might be reasonably used for the indirect calculation of the rate of disease per manifestation. On the other hand, if this type of careful monitoring is to be established for acquiring data, the diagnostician might just as

easily, and perhaps more accurately, assemble the direct information about the rate of disease per manifestation, and avoid the indirect Bayesian calculations.

3. Statistical Deficiencies

Regardless of the reliability of the numerical data, and regardless of whether the probabilities are calculated by direct or indirect methods, the cited statistical mechanisms of reasoning have several insurmountable defects as guidelines to clinical diagnosis.

a. The rate of error. In using rates of probability as a basis for decision, a statistician tries to minimize his risk of being wrong, but he is constantly willing to accept a certain rate of error in his deliberations. For example, a P value of 0.01, which is usually regarded as a high level of "statistical significance" for inferential decisions, means essentially that the statistician is willing to be wrong once in every 100 times that he makes the decision. At a P value of 0.05, which is the customary level of "statistical significance," the statistician is willing to be wrong once in every 20 times.

The rate of error that is acceptable to a statistician is intolerably high for doctors and patients. Patients would have no confidence in doctors, and doctors would have no confidence in their own work, if the doctor were blithely willing to be wrong in as many as 5% of the diverse decisions made during clinical reasoning. A thoughtful clinician recognizes that his reasoning may sometimes be incorrect and that his rate of error may often be higher than 5%, but he does not deliberately begin his rational activities with a complacent acceptance of any rate of error. In an era of science, the diagnostician wants to be exactly right, rather than approximately; and he is interested in scientific specificity and precision, rather than in statistical arithmetic and probabilities.

b. Atypical clinical situations. Just as a good computer program is evaluated by the way it manages exceptional rather than customary situations, a good scheme of diagnostic strategy should also be geared to management of the unusual or atypical. By this criterion of assessment, no existing statistical strategy can receive a passing grade. The statistical procedures are not prepared to cope with at least three of the most common types of "uncommon" event in diagnosis: the asymptomatic disease, the disease with bizarre complications, and the patient with multiple diseases.

As modern clinicians perform such screening tests as electrocardiography, chest X-ray, urinalysis, and cervical "Pap" smears to detect disease in asymptomatic patients, the incentive for each test is surely not a calculation of probabilities that a wholly asymptomatic patient has heart disease, tuberculosis, diabetes mellitus, or cervical cancer. Even if reliable statistical data could be collected about the prevalence of these diseases in asymptomatic patients, the diagnoses would still rest on the results of the tests, not on the arithmetical probabilities.

A separate problem is created by occasional patients with unusual or bizarre complications. Suppose a patient with an intestinally asymptomatic carcinoma of the rectum develops metastases to the pericardium, causing a large effusion and the presenting symptom of dyspnea due to pericardial tamponade. From the physical examination, chest X-ray, and paracentesis of pericardial fluid, the clinical diagnostician could readily infer the diagnosis. The statistical diagnostician would be left to explore numerical data, if any exist, about the probability of cancer of the rectum in patients with dyspnea.

Finally, the frequent use of a plethora of technologic tests has led to the realization that polypathic patients (with multiple diseases) frequently occur in modern medicine. The current statistical procedures make no provision for such patients. Thus, according to the statistical strategies, a patient with rectal bleeding might have either hemorrhoids or a cancer of the rectum or diverticula—but he could not have (as some patients do) two of these diseases or all three.

c. The absence of correlated explanation. As noted earlier, a critical role of diagnostic reasoning is not merely to identify diseases, but to explain manifestations. If the clinician does not follow the scientific pathway of correlating each "immediate cause" with its effect, he runs the risk of arriving at the right answers to the wrong questions. His conclusions may be accurate but impertinent (21), giving him results that are statistically satisfactory but scientifically defective, diagnostically inadequate, and therapeutically hazardous. Since statistical approaches in diagnosis are calculated to identify but not to correlate or to explain, they will inevitably reduce a clinician's effectiveness in meeting his daily challenges in patient care.

d. Clinical specificity vs general probability. Another major deficiency of current statistical approaches to diagnosis arises not from the validity of Bayesian inference in the calculations, not from the reliability of the basic data, and not from any of the features that have just been cited, but from the statistical necessity for using data that are often inappropriate for the calculations.

Let us assume, for the patient cited in earlier examples, that we had obtained reliable data about the occurrence rates of rectosigmoid cancer and of a suprapubic mass in the general population. The assumption is seldom valid, of course, since doctors can rarely, if ever, arrange for everyone in a population to be examined with procedures that are suitably applied and standardized. Nevertheless, for the sake of continuing the discussion, let us assume that the basic statistical data are reliable, and that the rate of rectosigmoid cancer in the general population is truly 1/120,000, and that the rate of suprapubic masses is truly 1/100,000. Let us further assume that the rate of a suprapubic mass in patients with rectosigmoid cancer has also been reliably determined and found to be 1/10. Let us finally assume that the probability of rectosigmoid cancer per suprapubic mass has been calculated by Bayesian methods or obtained directly, and that this value is truly 1/12.

Are these data, although reliable, really pertinent for the diagnostician's decision? Is he concerned with the statistical rate of events in either a general population or a diseased population—or is he trying to evaluate the phenomena observed in a particular patient? If the patient with the suprapubic mass were a 47-year-old white woman who had had diabetes mellitus since childhood, and who currently had a creamy vaginal discharge, would the clinician want to have data about the rate of rectosigmoid cancer in the general population, or would the clinician want the denominator population to consist of 47-year-old female juvenile diabetics with creamy vaginal discharges? Would he want to know about the rates of suprapubic mass in all people with rectosigmoid cancer, or in rectosigmoidally carcinomatous patients who also have the same demographic and clinical attributes as the patient at hand?

If the more specific types of statistical data were available in numbers large enough to be meaningful, the clinician would obviously prefer to use such data. But such specific data are generally not available or when available, not abundant—and so the imprecision of general statistics may be deliberately accepted

in order to make use of the quantification. To this indictment, a thoughtful statistician might reply that a scientific insistence on specificity would prevent statistics from ever being used in human reasoning, because a pedantic scientist could always find a descriptive attribute that was omitted from the assembled statistics. This issue of specificity thus leads to the core problem in the use of statistical data in clinical medicine.

Since diagnostic reasoning is always concerned with a particular patient, and since no two people are ever exactly the same, can any collection of statistics ever be applied to the individuality of a particular person? An antistatistical nihilist would probably answer this last question with a vigorous "No." The same answer would probably also be given, with or without antistatistical bias, by an "old school" clinician who resolutely believes that diagnostic judgment is, now and forever, an undefinable and indescribable act of intuitive artistry.

This rejection of statistics is probably warranted if we assume that statistics will continue to be used in the unscientific manner portrayed in the previous examples. On the other hand, if suitable specificity can be brought into both the statistical data and the clinical reasoning, a scientific clinician of the future can use statistics as a source of enormous illumination and assistance. To achieve this goal, however, requires a more effective recognition than has hitherto been given to the different types of decisions made by clinicians, and to the types of statistics needed for the decisions.

4. The Targets of Decision

The true value of statistics in clinicial reasoning has been obscured by a failure to distinguish the two different targets—intellectual and managerial—at which the decisions are aimed (22). The choice of a diagnostic name is an intellectual decision; the choice of a diagnostic test or of a mode of treatment is a managerial decison. In an intellectual decision, the final result is an identification or evaluation of an observed entity; in a managerial decision, the final result is a plan of action.

To select a mode of management for a patient, a clinician must rely on previous statistics; the only way in which he can customarily justify his decision is from a careful analysis of the results obtained in previous patients. He may be able to enumerate those results quantitatively, or he may document them only with recollections of "past experience," but the assessment is essentially a statistical procedure (3). In the intellectual choice of a diagnostic name, however, the clinician need not rely on statistical data. Regardless of what happened in previous patients, he can often justify his decision from the results obtained in a well-chosen paraclinical test. Thus, to select treatment for a patient with hemoptysis and lung cancer, the clinician must contemplate the statistical results of previous treatment in such patients; but in choices of diagnosis, if the bronchoscopic biopsy reveals an epidermoid carcinoma, a patient with hemoptysis has been shown to have lung cancer, regardless of the diagnoses suggested either by unquantified "past experience" or by statistical enumerations.

The cardinal defect of current statistical approaches to diagnosis, therefore, is not contained in any of the flaws that have already been discussed. The basic problem is that statistical tactics are being used for the wrong type of decision. Statistical data provide a necessary and vital background for managerial decisions in clinical science, but not for intellectual decisions. If the intellectual decision is the type of appraisal that produces such terms of evaluation as "good," "early," "improv-

ing," and "unsuccessful," the decision depends not on statistics, but on the specific criteria used for the appraisals (23). If the result of the intellectual decision is a diagnostic identification, then the decision depends on specific paraclinical evidence of the disease, or on eclectic criteria for establishing a diagnosis from various combinations of evidence (3, 23). If a diagnosis can not be conclusively established, however, the clinician's next move is toward an action decision, in which he either seeks additional information that will produce a conclusive diagnosis, or he plans a mode of therapy that seems optimal for managing the uncertainty.

Statistical approaches to diagnosis cannot be clinically satisfactory, therefore, because the data and the calculations used in the statistics are not appropriate for scientific precision in identifying disease, and are not directed at a managerial decision in planning either treatment or a further diagnostic work-up. When scientific precision is obtainable from a paraclinical test, the clinician will order that test whenever possible, and will heed its biologic information more than any statistical probabilities. When a precise diagnostic conclusion is not obtainable from the reasoning and the tests, the clinician will want to review statistics not about diagnosis, but about further management of the patient.

Many of the current statistical strategies in diagnosis are based on the naive belief that clinical activities have remained unchanged during the past century, and that the diagnostic process today is essentially the same as the "physical diagnosis" of the past, with clinical symptoms and signs as the main evidence that is converted into inferences about morphologic diseases. In that old fashion of "physical diagnosis," the clinician had very few procedures of paraclinical examination, and had to make his morphologic inferences as best he could, essentially devoid of ancillary tests. Such a clinician would gladly have welcomed any help he could get from Bayesian or other statistical strategies that might assist the intellectual transfer from clinical manifestations to morphologic diagnoses.

But this type of purely clinical "physical diagnosis" has been drastically altered by the availability of the roentgen ray, the biopsy, the endoscope, the myriads of laboratory tests, the surgical explorations, and all the other technologic procedures of modern medicine. In living patients, these procedures can identify morphologic diseases with precise specificity, and can demonstrate the many functional or chemical ailments that cannot be shown morphologically. Concomitant with all these changes in diagnosis, other advances in technology have provided a powerful therapeutic armamentarium that has drastically altered the types of management available as "action" to follow the diagnostic decisions of the past.

If statistics are to be helpful to a clinician who works with the technology of modern clinical science, the help will surely not come from applying new forms of statistical reasoning to archaic forms of clinical practice. Since the inherent validity of statistics is greatest for managerial rather than for intellectual decisions, the most potent role of statistics in diagnostic medicine today is not in the intellectual identification of disease, but in the managerial choice of diagnostic tests. With so much of the contemporary diagnostic process devoted to the "work-up," rather than to the reasoning, the selection of suitable tests has become a major new challenge for the clinician, and a major new hazard for the patient.

Many of the tests are dangerous, uncomfortable, and expensive (24). Many are superfluous, redundant, or inaccurate (25). They have been spawned promiscuously, without intensive regulation and often without suitable assessment of their worth, and they are often ordered thoughtlessly, without their results being sub-

jected to careful evaluation. To improve this lamentable situation, diagnosticians today are in desperate need of statistical assistance not for performing sterile calculations about the names of diseases, but for choosing the tests wisely, and for planning the diagnostic work-up effectively. A diagnostician needs statistical guidance to determine, at each stage of the diagnostic work-up, the costs, the risks, and the value of each diagnostic procedure.

For this guidance, a clinician could cogently and gratefully make use of appropriate statistical data. But no such data are available. The statistical approach to diagnostic reasoning has been devoted only to identifying "output" diseases from "input" manifestations. And the stages of the work-up have been delineated neither by clinicians nor by statisticians, so that the appropriate data cannot be obtained even if investigators had the inquiring incentive and suitable facilities. We cannot determine what are the values and hazards of ordering a test at one station in the reasoning, in contrast to another station, because the stations have not been delineated and correlated with the results of the tests. We cannot determine whether a group of tests should be ordered in sequence or as a single "battery," because the clinical and paraclinical constituents of a diagnostic sequence have not been stipulated.

If current statistical strategies in diagnosis are unacceptable, and if the stages of reasoning in the diagnostic work-up require a branching delineation of logical stations and decisions, then a clinician who wants to achieve effective diagnosis and effective statistics in modern medicine must arrange to specify the ingredients, the sequence, and the reasoning of those decisions. The discussions thus far have been concerned with the intellectual contents and intermediate strategies of the decisions. The next paper, which concludes this series, will deal with the algorithms, flow-charts, decision tables, and other tactics that can be used to specify those procedures.

REFERENCES

1. Feinstein, A. R. An analysis of diagnostic reasoning. I. The domains and disorders of clinical macrobiology. *Yale J. Biol. Med.* **46,** 212–232 (1973).
2. Feinstein, A. R. (Editorial) The basic elements of clinical science. *J. Chron. Dis.* **16,** 1125–1133 (1963).
3. Feinstein, A. R. "Clinical Judgment." Williams & Wilkins, Baltimore, 1967.
4. Mainland, D. "Elementary Medical Statistics." 2nd ed. Saunders, Philadelphia, 1963.
5. Murphy, E. A., and Abbey, H. The normal range—a common misuse. *J. Chron. Dis.* **20,** 79–88 (1967).
6. Herrera, L. The precision of percentiles in establishing normal limits in medicine. *J. Lab. Clin. Med.* **52,** 34 (1958).
7. Elveback, L. R., Guillier, C. L., and Keating, F. R. Health, normality and the ghost of Gauss. *J. Amer. Med. Ass.* **211,** 69–75 (1970).
8. Bernard, C. "An Introduction to the Study of Experimental Medicine." (Trans. by H. C. Greene.), p. 101. Collier, New York, 1961.
9. Feinstein, A. R., and Niebyl, J. R. Changes in the diagnostic process during 40 years of clinico-pathologic conferences. *Arch. Intern. Med.* **128,** 774–780 (1971).
10. Feinstein, A. R. The pre-therapeutic classification of co-morbidity in chronic disease. *J. Chron. Dis.* **23,** 455–469 (1970).
11. Durham, R. H. "Encyclopedia of Medical Syndromes." Hoeber Medical Division, Harper & Row, New York, 1960.
12. Himsworth, H. P. The syndrome of diabetes mellitus and its causes. *Lancet* **1,** 465–472 (1949).
13. London, S. J. The whimsy syndromes. The fine art of literary nosology. *Arch. Intern. Med.* **122,** 448–452 (1968).

14. Feinstein, A. R. Clinical epidemiology: II. The identification rates of disease. *Ann. Intern. Med.* **69,** 1037–1061 (1968).
15. Hogben, L. "Statistical Theory: The Relationship of Probability, Credibility and Error." Allen & Unwin, London, 1957.
16. Hartley, H. O. In Dr. Bayes' consulting room. *Amer. Statist.* **17,** 22–24 (1963).
17. Pearson, E. S. Discussion of Professor Smith's paper. *J. Roy. Statist. Soc. A* **128,** 490–492 (1965).
18. Bross, I. D. J. Applications of probability: Science vs. pseudoscience. *J. Amer. Statist. Ass.* **64,** 51–57 (1969).
19. Bruce, R. A., and Yarnall, S. R. Computer-aided diagnosis of cardiovascular disorders. *J. Chron. Dis.* **19,** 473–484 (1966).
20. Warner, H. A. A mathematical approach to medical diagnosis. Application to congenital heart disease. *J. Amer. Med. Ass.* **177,** 177–183 (1961).
21. Smith, V. M. The impertinent diagnosis. *Amer. J. Gastroenterol.* **48,** 334–342 (1967).
22. Feinstein, A. R. What kind of basic science for clinical medicine? *N. Engl. J. Med.* **283,** 847–852 (1970).
23. Feinstein, A. R. Taxonorics: I. Formulation of criteria. *Arch. Intern. Med.* **126,** 679–693 (1970).
24. Schimmel, E. The hazards of hospitalization. *Ann. Intern. Med.* **60,** 100–110 (1964).
25. Barnett, R. N. "Clinical Laboratory Statistics." Little, Brown, Boston, 1971.